## Summary

This is my document for my continuous assessment for Interactive Media Design and Visualisation. This document will outline how I came up with my dashboard visualisation and the steps in how I arrived at them. This continuous assessment was elapsed over three weeks. In my opinion, I felt the hardest part of the continuous assessment was gaining the data I wanted to use. I searched for a few hours on multiple occasions to find the datasets that suited me. Finally, I found it at <http://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=MUM01> . This gave me all the data I needed for multiple months in a half decent format which helped me to clean it and use it for myself in my project. In this project, I wanted to show the correlation between unemployment and the recession which occurred in 2008. I was trying to show the unemployment epidemic happening in Ireland at this current moment in time compared to before and after the Celtic Tiger era which impacted this country greatly.

## Background

I became interested in this topic when I was in secondary school a few years ago. I had an English teacher who would talk about current situations in the country and the recession seemed to come up a lot in discussion. At first I was not so interested but towards the end of my time in secondary school I began to see and feel the impact of the recession with family members and family friends losing jobs and homes first hand. I realised that everything my teacher had discussed with us in class was coming through. At first it was slowly, but month after month it gathered momentum.

This is what gave me the idea for this project. I spent weeks looking for the right dataset which I will go into further into this document. The aim of this project is to see just how long after events such as recessions and crashes are the effects felt by citizens within Ireland. I would like to see what other epidemics such as homelessness are effects due to these recessions and crashes and to see if there is a strong relationship between these two tragic social plights.

My continuous assessment will be different as I will be using this report and my visualisations to convey a powerful message to the reader. I want the reader to be shocked when they see my visualisations as they will instantly see how big the problem is in this country.

## Acquire

My research began looking at worldwide unemployment rate datasets which I accessed through some sites such as Eurostats.com and data.worldbank.org. From doing research on this I realised that I was finding data on such a large scale and I thought completing the project on something closer to home would be of much more benefit to the public which I had a stronger connection with from the same land. This is the reason why I chose my home nation, Ireland. I felt the terrible inflictions of the recession from 2008 first hand, as did several of my friends and family. This is the reason why I thought starting my research in this direction would be more interesting and of more benefit to me as I would see trends which I have a stronger bond with rather than the world which is a big place and I do not have direct connections to in most cases than not.

I began researching and looking for datasets on unemployment rates within Ireland. At first I found a considerable number of datasets which I was not worried about. I looked at these sets and thought that they could be used for comparisons more so down the road. I saved these datasets and kept moving on towards the data I needed.

After days of frustration looking for the right dataset I stumbled across the Central Statistics Office website and data portal. I searched through all categories getting datasets which were of no use to me and my goals that I set out at the start which I was determined to prove. Luckily I didn’t give up and I stumbled across the source I found from the Google search within the Central Statistics Office website which was great for my subject. The dataset was of the unemployment rate in Ireland from the year 1998 – 2016, month by month, and by genders. This was exactly what I was looking for. It included male and female information also which I will use to create some great visualisations for easy comparisons.

However, reviewing the data within the excel spreadsheet which the Central Statistic had put together was a mess. Trying to make sense of it was hard at first but from doing this a few times I gained a better grasp and understanding of the data that was on hand to me. Each time I learned more and more about the dataset I wanted to work with for this project. I thought this would be my number one dataset while completing this project.

## Parsing

At this point in time I was happy with the dataset I had gained. This would be the backbone of my project. The next stage I needed to do was to clean it and get it ready to be imported into Tableau. The dataset was a mess and it would take a lot of time and effort to clean this data manually. To combat against this massive time consumption totally manual process I decided to write a script in C#. This script took about an hour to perfect and I enjoyed doing it at the same time. In the dataset, the month and year was in one string value e.g. suppose we had January 2016, this was shown as “2016M01” which was not very useful to us in this format. My first step was to split these into a more usable format for Tableau. I decided to split the string into month and year columns in my new excel spreadsheet. When I ran the script, it produced an output for the months and years 225 times which was the exact number of times I needed them. Next all I had to do is to copy this data into the excel spreadsheet. I had no problems importing this, the columns were inserted perfect.

One problem with parsing the data was trying to complete the data with age groups all along the spreadsheet. If I dragged the cell downwards it would go down in jumps of years. This was not what I wanted. I wrote this extra piece of functionality into my script to produce the age group I needed and how many times it was needed. Next I copied this output from the console and it inserted with no problems again.

A very important piece of data was the last part to go into the spreadsheet. This was the number of Irish citizens who were unemployed. If I was to copy this information from the acquire phase to the parse phase it would be inserted into one cell as a long string value. This is not the expected outcome I was hoping for. I needed to prevent this from happening as all my previous work would have been a waste of time and I would have to enter the data one cell at a time. Luckily I found a solution. I opened a Microsoft word document, copied the text I needed, and converted it into a table. It was a simple approach for a big win on this project. It saved me an enormous amount of time if I had to go another route. I repeated this approach for another row in the acquire stage. My dataset was now taking the shape I wanted and could move to the next stage.

## Filter

This stage involved me removing data I found irrelevant for my data visualisations which will come. In this dataset, I only had to remove one column. This was the percentage of people currently unemployed. I thought Ii would allow Tableau to calculate the percentages for me as part of my data mining step.

## Data Mining

This step was carried out within Tableau to show some more complex tasks within the software. I looked at certain data by adding extra table calculations such as the average and sums to name a few options. This allowed me to see the entire sum of people unemployed within any given month.

## Represent

This stage involves me visualising the data which I have cleaned up and imported into Tableau for further analytic work. This was the most interesting part of the project as I seen trends that I did not notice before I started inserting the data into charts.

My first visualisation is something which opened my eyes to this problem. It is a horizontal bar chart with the number of unemployed people on the X-Axis and years on the Y-Axis. From this I can see a lot of detail already from a simple chart. From investigating this chart, I see that there was a massive rise in unemployment from 2008 – 2015. This began getting increasingly worse as the year 2011 where there was a big jump in the unemployment rate. From glancing at the chart, I see that this year was the worst year in the data I have for unemployment. I would like to investigate this more and see what other than the recession could have effected this big change.

My second visualisation revealed another unknown pattern to me which I had no prior knowledge about. From looking at this chart I can see that there is a big difference in the age groups and genders. In saying this, I can reveal that there is not a big gap between 15 – 24-year-old males and females in relation to unemployment. However, if we look at the next group and gender i.e. the 25 – 74 - year-olds there is a great difference in unemployment. It seems to me that males have a higher rate of unemployment than their female counter parts. This is an odd pattern and I would like to know why there is a greater number of males unemployed than women. Another great feature of Tableau is the use of being able to forecast predictions for the coming years. I predicted the next three years of unemployment and on each occasion, it is showing that unemployment is rising towards levels like that of 2014.

## Refine

To improve my first visualisation I have added in more colour. The colour code is low unemployment is green as the lower that rate is the better and red for higher unemployment as that is bad. There are also different shades within the charts to show the transition between lowness and highness with regards to unemployment. The user can also click on whatever bar they want and they are provided with extra information about the bar they have chosen.

To improve the second visualisation I have added labels and used the mean average to show the bar chart. The labels represent the mean average for the years on the X-Axis. The user can hover over which ever bar they choose and they are show extra information, like the first visualisation.

## Interact

In visualisation 1, the user can click on the bar which they want to see more information about. Once clicked a tool tip is presented to them This shows some more information about this bar to the user.

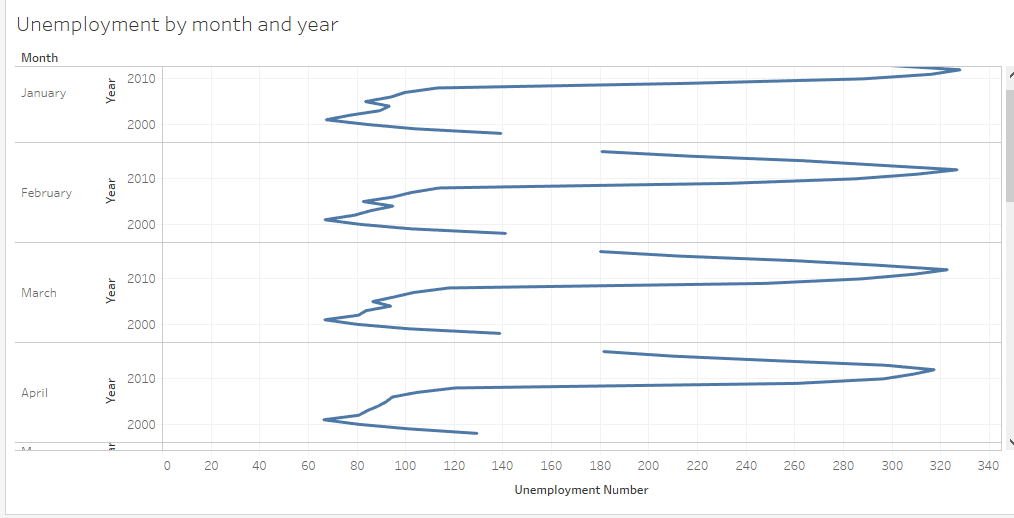
To manipulate the data in visualisation 2, the user is given a menu on the right-hand side of the screen. On this side of the screen the user can click on the group of values which ever they wish to view. By doing this, the user is shown the corresponding values only. This gives the user some more power over the visualisation as they can see whatever they wish to see.

## Problems and solutions

At the start of the project, getting the data into the format that would allow me import it into Tableau provided a great challenge to me. The solution to this problem was writing a script in C# which would output the chosen text into the format that I needed to be able to import the excel spreadsheet into Tableau. This was splitting out the months and the years into two separate string values.

Likewise, with the numeric values. When I was copying the long string values into the parse stage from the acquire stage I had problems getting them to fill the columns which corresponded to their values. When I would paste them over they would go into one very long cell which is not what I wanted. The solution to this problem was by copying the chosen text into a Microsoft word document and converting that text into a table. This allowed me then to copy the whole table into my excel spreadsheet the way I needed it to be.

Some of my visualisations did not work in conveying the message I wish to show to the users. One chart which completely failed on me was the line chart. This chart is below:



There is a lot of information on this chart which would confuse the user. The year starts at 2000 on the Y-Axis and as the axis goes up the year increases. This is not visible to the user. There is also 12 separate charts on this which is a lot of information to the user.

## Conclusion